

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Mathews, Lawrence S.  
Vale, Wylie W.  
Tsuchida, Kunihiro
- (ii) TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF  
RECEPTOR(S) OF THE ACTIVIN/TGF-BETA SUPERFAMILY
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
(B) STREET: 444 South Flower Street, Suite 2000  
(C) CITY: Los Angeles  
(D) STATE: CA  
(E) COUNTRY: USA  
(F) ZIP: 90071
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 08/300,584  
(B) FILING DATE: 02-SEP-1994
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 07/880,220  
(B) FILING DATE: 08-MAY-1992
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 07/773,229  
(B) FILING DATE: 09-OCT-1991
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 07/698,709  
(B) FILING DATE: 10-MAY-1991
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Reiter, Stephen E.  
(B) REGISTRATION NUMBER: 31,192  
(C) REFERENCE/DOCKET NUMBER: P41 9927
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 619-546-4737  
(B) TELEFAX: 619-546-9392

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2563 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

00742604-121900

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 71..1609

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCCGAGGAA GACCCAGGGA ACTGGATATC TAGCGAGAAC TTCCTACGGC TTCTCCGGCG	60
CCTCGGGAAA ATG GGA GCT GCT GCA AAG TTG GCG TTC GCC GTC TTT CTT	109
Met Gly Ala Ala Ala Lys Leu Ala Phe Ala Val Phe Leu	
1 5 10	
ATC TCT TGC TCT TCA GGT GCT ATA CTT GGC AGA TCA GAA ACT CAG GAG	157
Ile Ser Cys Ser Ser Gly Ala Ile Leu Gly Arg Ser Glu Thr Gln Glu	
15 20 25	
TGT CTT TTC TTT AAT GCT AAT TGG GAA AGA GAC AGA ACC AAC CAG ACT	205
Cys Leu Phe Phe Asn Ala Asn Trp Glu Arg Asp Arg Thr Asn Gln Thr	
30 35 40 45	
GGT GTT GAA CCT TGC TAT GGT GAT AAA GAT AAA CGG CGA CAT TGT TTT	253
Gly Val Glu Pro Cys Tyr Gly Asp Lys Asp Lys Arg Arg His Cys Phe	
50 55 60	
GCT ACC TGG AAG AAT ATT TCT GGT TCC ATT GAA ATA GTG AAG CAA GGT	301
Ala Thr Trp Lys Asn Ile Ser Gly Ser Ile Glu Ile Val Lys Gln Gly	
65 70 75	
TGT TGG CTG GAT GAT ATC AAC TGC TAT GAC AGG ACT GAT TGT ATA GAA	349
Cys Trp Leu Asp Asp Ile Asn Cys Tyr Asp Arg Thr Asp Cys Ile Glu	
80 85 90	
AAA AAA GAC AGC CCT GAA GTG TAC TTT TGT TGC TGT GAG GGC AAT ATG	397
Lys Lys Asp Ser Pro Glu Val Tyr Phe Cys Cys Cys Glu Gly Asn Met	
95 100 105	
TGT AAT GAA AAG TTC TCT TAT TTT CCG GAG ATG GAA GTC ACA CAG CCC	445
Cys Asn Glu Lys Phe Ser Tyr Phe Pro Glu Met Glu Val Thr Gln Pro	
110 115 120 125	
ACT TCA AAT CCT GTT ACA CCG AAG CCA CCC TAT TAC AAC ATT CTG CTG	493
Thr Ser Asn Pro Val Thr Pro Lys Pro Pro Tyr Tyr Asn Ile Leu Leu	
130 135 140	
TAT TCC TTG GTA CCA CTA ATG TTA ATT GCA GGA ATT GTC ATT TGT GCA	541
Tyr Ser Leu Val Pro Leu Met Leu Ile Ala Gly Ile Val Ile Cys Ala	
145 150 155	
TTT TGG GTG TAC AGA CAT CAC AAG ATG GCC TAC CCT CCT GTA CTT GTT	589
Phe Trp Val Tyr Arg His His Lys Met Ala Tyr Pro Pro Val Leu Val	
160 165 170	
CCT ACT CAA GAC CCA GGA CCA CCC CCA CCT TCC CCA TTA CTA GGG TTG	637
Pro Thr Gln Asp Pro Gly Pro Pro Pro Pro Ser Pro Leu Leu Gly Leu	
175 180 185	
AAG CCA TTG CAG CTG TTA GAA GTG AAA GCA AGG GGA AGA TTT GGT TGT	685
Lys Pro Leu Gln Leu Leu Glu Val Lys Ala Arg Gly Arg Phe Gly Cys	
190 195 200 205	
GTC TGG AAA GCC CAG TTG CTC AAT GAA TAT GTG GCT GTC AAA ATA TTT	733
Val Trp Lys Ala Gln Leu Leu Asn Glu Tyr Val Ala Val Lys Ile Phe	
210 215 220	

005727-73927269

CCA	ATA	CAG	GAC	AAA	CAG	TCC	TGG	CAG	AAT	GAA	TAT	GAA	GTC	TAT	AGT	781
Pro	Ile	Gln	Asp	Lys	Gln	Ser	Trp	Gln	Asn	Glu	Tyr	Glu	Val	Tyr	Ser	
			225					230					235			
CTA	CCT	GGA	ATG	AAG	CAT	GAG	AAC	ATA	CTA	CAG	TTC	ATT	GGT	GCA	GAG	829
Leu	Pro	Gly	Met	Lys	His	Glu	Asn	Ile	Leu	Gln	Phe	Ile	Gly	Ala	Glu	
		240					245					250				
AAA	AGA	GGC	ACC	AGT	GTG	GAT	GTG	GAC	CTG	TGG	CTA	ATC	ACA	GCA	TTT	877
Lys	Arg	Gly	Thr	Ser	Val	Asp	Val	Asp	Leu	Trp	Leu	Ile	Thr	Ala	Phe	
	255					260					265					
CAT	GAA	AAG	GGC	TCA	CTG	TCA	GAC	TTT	CTT	AAG	GCT	AAT	GTG	GTC	TCT	925
His	Glu	Lys	Gly	Ser	Leu	Ser	Asp	Phe	Leu	Lys	Ala	Asn	Val	Val	Ser	
270					275					280					285	
TGG	AAT	GAA	CTT	TGT	CAT	ATT	GCA	GAA	ACC	ATG	GCT	AGA	GGA	TTG	GCA	973
Trp	Asn	Glu	Leu	Cys	His	Ile	Ala	Glu	Thr	Met	Ala	Arg	Gly	Leu	Ala	
				290					295					300		
TAT	TTA	CAT	GAG	GAT	ATA	CCT	GGC	TTA	AAA	GAT	GGC	CAC	AAG	CCT	GCA	1021
Tyr	Leu	His	Glu	Asp	Ile	Pro	Gly	Leu	Lys	Asp	Gly	His	Lys	Pro	Ala	
			305					310					315			
ATC	TCT	CAC	AGG	GAC	ATC	AAA	AGT	AAA	AAT	GTG	CTG	TTG	AAA	AAC	AAT	1069
Ile	Ser	His	Arg	Asp	Ile	Lys	Ser	Lys	Asn	Val	Leu	Leu	Lys	Asn	Asn	
		320					325					330				
CTG	ACA	GCT	TGC	ATT	GCT	GAC	TTT	GGG	TTG	GCC	TTA	AAG	TTC	GAG	GCT	1117
Leu	Thr	Ala	Cys	Ile	Ala	Asp	Phe	Gly	Leu	Ala	Leu	Lys	Phe	Glu	Ala	
	335					340					345					
GGC	AAG	TCT	GCA	GGT	GAC	ACC	CAT	GGG	CAG	GTT	GGT	ACC	CGG	AGG	TAT	1165
Gly	Lys	Ser	Ala	Gly	Asp	Thr	His	Gly	Gln	Val	Gly	Thr	Arg	Arg	Tyr	
350					355				360						365	
ATG	GCT	CCA	GAG	GTG	TTG	GAG	GGT	GCT	ATA	AAC	TTC	CAA	AGG	GAC	GCA	1213
Met	Ala	Pro	Glu	Val	Leu	Glu	Gly	Ala	Ile	Asn	Phe	Gln	Arg	Asp	Ala	
				370					375					380		
TTT	CTG	AGG	ATA	GAT	ATG	TAC	GCC	ATG	GGA	TTA	GTC	CTA	TGG	GAA	TTG	1261
Phe	Leu	Arg	Ile	Asp	Met	Tyr	Ala	Met	Gly	Leu	Val	Leu	Trp	Glu	Leu	
			385					390					395			
GCT	TCT	CGT	TGC	ACT	GCT	GCA	GAT	GGA	CCC	GTA	GAT	GAG	TAC	ATG	TTA	1309
Ala	Ser	Arg	Cys	Thr	Ala	Ala	Asp	Gly	Pro	Val	Asp	Glu	Tyr	Met	Leu	
		400					405					410				
CCA	TTT	GAG	GAA	GAA	ATT	GGC	CAG	CAT	CCA	TCT	CTT	GAA	GAT	ATG	CAG	1357
Pro	Phe	Glu	Glu	Glu	Ile	Gly	Gln	His	Pro	Ser	Leu	Glu	Asp	Met	Gln	
	415					420					425					
GAA	GTT	GTT	GTG	CAT	AAA	AAA	AAG	AGG	CCT	GTT	TTA	AGA	GAT	TAT	TGG	1405
Glu	Val	Val	Val	His	Lys	Lys	Lys	Arg	Pro	Val	Leu	Arg	Asp	Tyr	Trp	
430					435				440						445	
CAG	AAA	CAT	GCA	GGA	ATG	GCA	ATG	CTC	TGT	GAA	ACG	ATA	GAA	GAA	TGT	1453
Gln	Lys	His	Ala	Gly	Met	Ala	Met	Leu	Cys	Glu	Thr	Ile	Glu	Glu	Cys	
				450					455					460		
TGG	GAT	CAT	GAT	GCA	GAA	GCC	AGG	TTA	TCA	GCT	GGA	TGT	GTA	GGT	GAA	1501
Trp	Asp	His	Asp	Ala	Glu	Ala	Arg	Leu	Ser	Ala	Gly	Cys	Val	Gly	Glu	
			465				470						475			
AGA	ATT	ACT	CAG	ATG	CAA	AGA	CTA	ACA	AAT	ATC	ATT	ACT	ACA	GAG	GAC	1549
Arg	Ile	Thr	Gln	Met	Gln	Arg	Leu	Thr	Asn	Ile	Ile	Thr	Thr	Glu	Asp	
			480				485						490			

006TET 16927460

ATT GTA ACA GTG GTC ACA ATG GTG ACA AAT GTT GAC TTT CCT CCC AAA	1597
Ile Val Thr Val Val Thr Met Val Thr Asn Val Asp Phe Pro Pro Lys	
495 500 505	
GAA TCT AGT CTA TGATGGTGGC ACCGTCTGTA CACACTGAGG ACTGGGACTC	1649
Glu Ser Ser Leu	
510	
TGAACTGGAG CTGCTAAGCT AAGGAAAGTG CTTAGTTGAT TTTCTGTGTG AAATGAGTAG	1709
GATGCCCTCCA GGACATGTAC GCAAGCAGCC CCTTGTGGAA AGCATGGATC TGGGAGATGG	1769
ATCTGGGAAA CTTACTGCAT CGTCTGCAGC ACAGATATGA AGAGGAGTCT AAGGGAAAAG	1829
CTGCAAACTG TAAAGAACTT CTGAAAATGT ACTCGAAGAA TGTGGCCCTC TCCAAATCAA	1889
GGATCTTTTG GACCTGGCTA ATCAAGTATT TGCAAACTG ACATCAGATT TCTTAATGTC	1949
TGTCAGAAGA CACTAATTCC TTAAATGAAC TACTGCTATT TTTTAAAT GAAAACTTT	2009
TCATTTTCAGA TTTTAAAAG GGTAACCTTT TATTGCATTT GCTGTTGTTT CTATAAATGA	2069
CTATTGTAAT GCCAACATGA CACAGCTTGT GAATGTGTAG TGTGCTGCTG TTCTGTGTAC	2129
ATAGTCATCA AAGTGGGGTA CAGTAAAGAG GCTTCCAAGC ATTACTTTAA CCTCCCTCAA	2189
CAAGGTATAC CTCAGTTCCA CGGTTGTAA ATTATAAAAT TGAAAACACT AACAGAATTT	2249
GAATAAATCA GTCCATGTTT TATAACAAGG TTAATTACAA ATTCACTGTG TTATTTAAGA	2309
AAAAATGGTA AGCTATGCTT AGTGCCAATA GTAAGTGGCT ATTTGTAAAG CAGTGTTTTA	2369
GCTTTTCTTC TACTGGCTTG TAATTTAGGG AAAACAAGTG CTGTCTTTGA AATGGAAAAG	2429
AATATGGTGT CACCCTACCC CCCATACTTA TATCAAGGTC CCAAAATATT CTTTCCATT	2489
TCAAAGACAG CACTTTGAAA ACCCTAAATT ACAAGCCAGT AGAAGAAAAG CTAAAACACG	2549
CTTTACAAAT AGCC	2563

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 513 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Gly	Ala	Ala	Ala	Lys	Leu	Ala	Phe	Ala	Val	Phe	Leu	Ile	Ser	Cys
1				5					10					15	
Ser	Ser	Gly	Ala	Ile	Leu	Gly	Arg	Ser	Glu	Thr	Gln	Glu	Cys	Leu	Phe
		20						25					30		
Phe	Asn	Ala	Asn	Trp	Glu	Arg	Asp	Arg	Thr	Asn	Gln	Thr	Gly	Val	Glu
		35				40						45			
Pro	Cys	Tyr	Gly	Asp	Lys	Asp	Lys	Arg	Arg	His	Cys	Phe	Ala	Thr	Trp
	50				55						60				
Lys	Asn	Ile	Ser	Gly	Ser	Ile	Glu	Ile	Val	Lys	Gln	Gly	Cys	Trp	Leu
65					70					75					80

Asp	Asp	Ile	Asn	Cys 85	Tyr	Asp	Arg	Thr	Asp 90	Cys	Ile	Glu	Lys	Lys 95	Asp
Ser	Pro	Glu	Val 100	Tyr	Phe	Cys	Cys	Cys 105	Glu	Gly	Asn	Met	Cys 110	Asn	Glu
Lys	Phe	Ser 115	Tyr	Phe	Pro	Glu	Met 120	Glu	Val	Thr	Gln	Pro 125	Thr	Ser	Asn
Pro	Val 130	Thr	Pro	Lys	Pro	Pro 135	Tyr	Tyr	Asn	Ile	Leu 140	Leu	Tyr	Ser	Leu
Val 145	Pro	Leu	Met	Leu	Ile 150	Ala	Gly	Ile	Val	Ile 155	Cys	Ala	Phe	Trp	Val 160
Tyr	Arg	His	His	Lys 165	Met	Ala	Tyr	Pro	Pro 170	Val	Leu	Val	Pro	Thr 175	Gln
Asp	Pro	Gly	Pro 180	Pro	Pro	Pro	Ser	Pro 185	Leu	Leu	Gly	Leu	Lys 190	Pro	Leu
Gln	Leu	Leu 195	Glu	Val	Lys	Ala	Arg 200	Gly	Arg	Phe	Gly	Cys 205	Val	Trp	Lys
Ala	Gln 210	Leu	Leu	Asn	Glu	Tyr 215	Val	Ala	Val	Lys	Ile 220	Phe	Pro	Ile	Gln
Asp 225	Lys	Gln	Ser	Trp	Gln 230	Asn	Glu	Tyr	Glu	Val 235	Tyr	Ser	Leu	Pro	Gly 240
Met	Lys	His	Glu	Asn 245	Ile	Leu	Gln	Phe	Ile 250	Gly	Ala	Glu	Lys	Arg 255	Gly
Thr	Ser	Val	Asp 260	Val	Asp	Leu	Trp	Leu 265	Ile	Thr	Ala	Phe	His 270	Glu	Lys
Gly	Ser	Leu 275	Ser	Asp	Phe	Leu	Lys 280	Ala	Asn	Val	Val	Ser 285	Trp	Asn	Glu
Leu	Cys 290	His	Ile	Ala	Glu	Thr 295	Met	Ala	Arg	Gly	Leu 300	Ala	Tyr	Leu	His
Glu 305	Asp	Ile	Pro	Gly	Leu 310	Lys	Asp	Gly	His	Lys 315	Pro	Ala	Ile	Ser	His 320
Arg	Asp	Ile	Lys	Ser 325	Lys	Asn	Val	Leu	Leu 330	Lys	Asn	Asn	Leu	Thr 335	Ala
Cys	Ile	Ala	Asp 340	Phe	Gly	Leu	Ala	Leu 345	Lys	Phe	Glu	Ala	Gly 350	Lys	Ser
Ala	Gly	Asp 355	Thr	His	Gly	Gln	Val 360	Gly	Thr	Arg	Arg	Tyr 365	Met	Ala	Pro
Glu	Val 370	Leu	Glu	Gly	Ala	Ile 375	Asn	Phe	Gln	Arg	Asp 380	Ala	Phe	Leu	Arg
Ile 385	Asp	Met	Tyr	Ala	Met 390	Gly	Leu	Val	Leu	Trp 395	Glu	Leu	Ala	Ser	Arg 400
Cys	Thr	Ala	Ala	Asp 405	Gly	Pro	Val	Asp	Glu 410	Tyr	Met	Leu	Pro	Phe 415	Glu
Glu	Glu	Ile	Gly 420	Gln	His	Pro	Ser	Leu 425	Glu	Asp	Met	Gln	Glu 430	Val	Val

Leu

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2335 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:  
(B) CLONE: XACTR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGCCCCACAC	AGTGCACTGA	ATAATAGCCG	GTGCGGCCCC	TCCCCCTCTTT	CCCTGGCAGT	60
TGTGTATCTG	TCACATTGAA	GTTTGGGCTC	CTGTGAGTCT	GAGCCTCCCC	CTGTGTCTCA	120
TGTGAAGCTG	CTGCTGCAGA	AGGTGGAGTC	GTTGCATGAG	GGTGGGGGGA	GTCTGCTGCTG	180
TTTGATCTGC	CTCTGCTCCC	CATTCACT	CTCATTTCAT	TCCCACGGAT	CCACATTACA	240
ACTCGCCTTT	AACCCCTTCC	CTGGCGGAGC	CCACGCGTCT	TTCATCCCTC	CTGCCGCGGC	300
CGCTGAGCGA	CCAGAGCGCG	ACATTGTTGC	GGCGGGGGAT	TGGGCGACAT	TGTTGCGAAT	360
AATCGGAGCT	GCTGGGGGGG	AACTGATACA	ACGTTGCGAC	TGTAAAGGAA	TTAACTCGGC	420
CGAATGGGAT	TTTATCTGTG	TCGGTGAGAG	AAGCGGATCC	CAGGAGC	ATG GGG GCG Met Gly Ala 1	476
TCT GTA GCG CTG ACT TTT CTA CTT CTT CTT GCA ACT TTC CGC GCA GGC						524
Ser Val 5 Ala Leu Thr Phe Leu 10 Leu Leu Leu Ala 15 Thr Phe Arg Ala Gly						
TCA GGA CAC GAT GAA GTG GAG ACA AGA GAG TGC ATC TAT TAC AAT GCC						572
Ser Gly His Asp Glu Val Glu Thr Arg Glu Cys Ile Tyr Tyr Asn Ala						
20 25 30 35						

AAC TGG GAA CTG GAG AAG ACC AAC CAA AGT GGG GTG GAA AGC TGC GAA Asn Trp Glu Leu Glu Lys Thr Asn Gln Ser Gly Val Glu Ser Cys Glu 40 45 50	620
GGG GAA AAG GAC AAG CGA CTC CAC TGT TAC GCG TCT TGG AGG AAC AAT Gly Glu Lys Asp Lys Arg Leu His Cys Tyr Ala Ser Trp Arg Asn Asn 55 60 65	668
TCG GGC TTC ATA GAG CTG GTG AAA AAA GGA TGC TGG CTG GAT GAC TTC Ser Gly Phe Ile Glu Leu Val Lys Lys Gly Cys Trp Leu Asp Asp Phe 70 75 80	716
AAC TGT TAT GAC AGA CAG GAA TGT ATT GCC AAG GAA GAA AAC CCC CAA Asn Cys Tyr Asp Arg Gln Glu Cys Ile Ala Lys Glu Glu Asn Pro Gln 85 90 95	764
GTC TTT TTC TGC TGC TGC GAG GGA AAC TAC TGC AAC AAG AAA TTT ACT Val Phe Phe Cys Cys Cys Glu Gly Asn Tyr Cys Asn Lys Lys Phe Thr 100 105 110 115	812
CAT TTG CCT GAA GTC GAA ACA TTT GAT CCG AAG CCC CAG CCG TCA GCC His Leu Pro Glu Val Glu Thr Phe Asp Pro Lys Pro Gln Pro Ser Ala 120 125 130	860
TCC GTA CTG AAC ATT CTG ATC TAT TCC CTG CTT CCA ATT GTT GGT CTT Ser Val Leu Asn Ile Leu Ile Tyr Ser Leu Leu Pro Ile Val Gly Leu 135 140 145	908
TCC ATG GCA ATT CTC CTG GCG TTC TGG ATG TAC CGT CAT CGA AAG CCT Ser Met Ala Ile Leu Leu Ala Phe Trp Met Tyr Arg His Arg Lys Pro 150 155 160	956
CCC TAC GGG CAT GTA GAG ATC AAT GAG GAC CCC GGT CTG CCC CCT CCA Pro Tyr Gly His Val Glu Ile Asn Glu Asp Pro Gly Leu Pro Pro Pro 165 170 175	1004
TCT CCT CTG GTC GGG CTG AAG CCG CTG CAG TTG CTG GAG ATA AAG GCG Ser Pro Leu Val Gly Leu Lys Pro Leu Gln Leu Leu Glu Ile Lys Ala 180 185 190 195	1052
CGA GGC CGT TTC GGT TGC GTC TGG AAA GCT CGT CTG CTG AAT GAA TAT Arg Gly Arg Phe Gly Cys Val Trp Lys Ala Arg Leu Leu Asn Glu Tyr 200 205 210	1100
GTC GCA GTG AAA ATC TTC CCC GTG CAG GAT AAG CAG TCG TGG CAG TGT Val Ala Val Lys Ile Phe Pro Val Gln Asp Lys Gln Ser Trp Gln Cys 215 220 225	1148
GAG AAA GAG ATC TTC ACC ACG CCG GGC ATG AAA CAT GAA AAC CTA TTG Glu Lys Glu Ile Phe Thr Thr Pro Gly Met Lys His Glu Asn Leu Leu 230 235 240	1196
GAG TTC ATT GCC GCT GAG AAG AGG GGA AGC AAC CTG GAG ATG GAG CTG Glu Phe Ile Ala Ala Glu Lys Arg Gly Ser Asn Leu Glu Met Glu Leu 245 250 255	1244
TGG CTC ATC ACT GCA TTT CAT GAT AAG GGT TCT CTG ACG GAC TAC CTG Trp Leu Ile Thr Ala Phe His Asp Lys Gly Ser Leu Thr Asp Tyr Leu 260 265 270 275	1292
AAA GGG AAC TTG GTG AGC TGG AAT GAA CTG TGT CAC ATA ACA GAA ACA Lys Gly Asn Leu Val Ser Trp Asn Glu Leu Cys His Ile Thr Glu Thr 280 285 290	1340
ATG GCT CGT GGG CTG GCC TAC TTA CAT GAA GAT GTG CCC CGC TGT AAA Met Ala Arg Gly Leu Ala Tyr Leu His Glu Asp Val Pro Arg Cys Lys 295 300 305	1388

00742584-12100





(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 510 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met 1	Gly	Ala	Ser	Val 5	Ala	Leu	Thr	Phe	Leu 10	Leu	Leu	Leu	Ala	Thr 15	Phe
Arg	Ala	Gly	Ser 20	Gly	His	Asp	Glu	Val 25	Glu	Thr	Arg	Glu	Cys 30	Ile	Tyr
Tyr	Asn 35	Ala	Asn	Trp	Glu	Leu	Glu 40	Lys	Thr	Asn	Gln	Ser 45	Gly	Val	Glu
Ser	Cys 50	Glu	Gly	Glu	Lys	Asp 55	Lys	Arg	Leu	His	Cys 60	Tyr	Ala	Ser	Trp
Arg 65	Asn	Asn	Ser	Gly	Phe 70	Ile	Glu	Leu	Val	Lys 75	Lys	Gly	Cys	Trp	Leu 80
Asp	Asp	Phe	Asn	Cys 85	Tyr	Asp	Arg	Gln	Glu 90	Cys	Ile	Ala	Lys	Glu 95	Glu
Asn	Pro	Gln	Val 100	Phe	Phe	Cys	Cys	Cys 105	Glu	Gly	Asn	Tyr	Cys 110	Asn	Lys
Lys	Phe	Thr 115	His	Leu	Pro	Glu	Val 120	Glu	Thr	Phe	Asp	Pro 125	Lys	Pro	Gln
Pro	Ser 130	Ala	Ser	Val	Leu	Asn 135	Ile	Leu	Ile	Tyr	Ser 140	Leu	Leu	Pro	Ile
Val 145	Gly	Leu	Ser	Met	Ala 150	Ile	Leu	Leu	Ala	Phe 155	Trp	Met	Tyr	Arg	His 160
Arg	Lys	Pro	Pro 165	Tyr	Gly	His	Val	Glu	Ile 170	Asn	Glu	Asp	Pro	Gly 175	Leu
Pro	Pro	Pro	Ser 180	Pro	Leu	Val	Gly	Leu 185	Lys	Pro	Leu	Gln	Leu 190	Leu	Glu
Ile	Lys	Ala 195	Arg	Gly	Arg	Phe	Gly 200	Cys	Val	Trp	Lys	Ala 205	Arg	Leu	Leu
Asn	Glu 210	Tyr	Val	Ala	Val	Lys 215	Ile	Phe	Pro	Val	Gln 220	Asp	Lys	Gln	Ser
Trp 225	Gln	Cys	Glu	Lys	Glu 230	Ile	Phe	Thr	Thr	Pro 235	Gly	Met	Lys	His	Glu 240
Asn	Leu	Leu	Glu	Phe 245	Ile	Ala	Ala	Glu	Lys 250	Arg	Gly	Ser	Asn	Leu 255	Glu
Met	Glu	Leu	Trp 260	Leu	Ile	Thr	Ala	Phe 265	His	Asp	Lys	Gly	Ser 270	Leu	Thr
Asp	Tyr	Leu 275	Lys	Gly	Asn	Leu	Val 280	Ser	Trp	Asn	Glu	Leu 285	Cys	His	Ile
Thr	Glu 290	Thr	Met	Ala	Arg	Gly 295	Leu	Ala	Tyr	Leu	His 300	Glu	Asp	Val	Pro

Arg Cys Lys Gly Glu Gly His Lys Pro Ala Ile Ala His Arg Asp Phe  
 305 310 315 320  
 Lys Ser Lys Asn Val Leu Leu Arg Asn Asp Leu Thr Ala Ile Leu Ala  
 325 330 335  
 Asp Phe Gly Leu Ala Val Arg Phe Glu Pro Gly Lys Pro Pro Gly Asp  
 340 345 350  
 Thr His Gly Gln Val Gly Thr Arg Arg Tyr Met Ala Pro Glu Val Leu  
 355 360 365  
 Glu Gly Ala Ile Asn Phe Gln Arg Asp Ser Phe Leu Arg Ile Asp Met  
 370 375 380  
 Tyr Ala Met Gly Leu Val Leu Trp Glu Ile Val Ser Arg Cys Thr Ala  
 385 390 395 400  
 Ala Asp Gly Pro Val Asp Glu Tyr Leu Leu Pro Phe Glu Glu Glu Ile  
 405 410 415  
 Gly Gln His Pro Ser Leu Glu Asp Leu Gln Glu Val Val Val His Lys  
 420 425 430  
 Lys Ile Arg Pro Val Phe Lys Asp His Trp Leu Lys His Pro Gly Leu  
 435 440 445  
 Ala Gln Leu Cys Val Thr Ile Glu Glu Cys Trp Asp His Asp Ala Glu  
 450 455 460  
 Ala Arg Leu Ser Ala Gly Cys Val Glu Glu Arg Ile Ser Gln Ile Arg  
 465 470 475 480  
 Lys Ser Val Asn Gly Thr Thr Ser Asp Cys Leu Val Ser Ile Val Thr  
 485 490 495  
 Ser Val Thr Asn Val Asp Leu Pro Pro Lys Glu Ser Ser Ile  
 500 505 510

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asp Leu Lys Pro Glu Asn  
 1 5

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

004466-1032460

(ix) FEATURE:  
(A) NAME/KEY: Modified-site  
(B) LOCATION: 2  
(D) OTHER INFORMATION: /note= "Xaa at position 2 is either  
"Thr" or "Ser"."

(ix) FEATURE:  
(A) NAME/KEY: Modified-site  
(B) LOCATION: 5  
(D) OTHER INFORMATION: /note= "Xaa a position 5 is either  
"Tyr" or "Phe"."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Gly Xaa Xaa Xaa Xaa Xaa  
1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Asp Ile Lys Ser Lys Asn  
1 5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gly Thr Arg Arg Tyr Met  
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

006727-43924260

Sub  
C1

Asp/Leu Ala Ala Arg Asn  
1 5

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: both

(v) ~~FRAGMENT~~ TYPE: internal

(A) NAME/KEY: Modified-site  
(B) LOCATION: 3  
(D) OTHER INFORMATION: /note= "Xaa at position 3 is either  
"Ile" or "Val"."

(A) NAME/KEY: Modified-site  
(B) LOCATION: 4  
(D) OTHER INFORMATION: /note= "Xaa at position 4 is either  
"Lys" or "Arg"."

(A) NAME/KEY: Modified-site  
(B) LOCATION: 6  
(D) OTHER INFORMATION: /note= "Xaa at position 6 is either  
"Thr" or "Met"."

Xaa Pro Xaa Xaa Trp Xaa  
1 5

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1602 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: both

```
(ix) FEATURE:
      (A) NAME/KEY: CDS
      (B) LOCATION: 72..1553
```

CCCCGGGAAC TTCAAAGCGC GCTGCGGCGG CGCTCTGGA CCCCAGGCC TTGCACCGCC  
GCGGGGTGGC C ATG ACC CCA GCG CGC CGC TCC GCA CTG AGC CTG GCC CTC  
Met Thr Pro Ala Arg Arg Ser Ala Leu Ser Leu Ala Leu  
1 5 10

60

110

10

[illegible]

Sub C'

ATA GTG ACC GTG GCT GGA ATG GTC AAG TTG GCG CTT TCA ATA GCG AGT Ile Val Thr Val Ala Gly Met Val Lys Leu Ala Leu Ser Ile Ala Ser	974
290 295 300	
GGT CTG GCT CAC CTA CAC ATG GAG ATC GTG GGC ACT CAA GGT AAG CCT Gly Leu Ala His Leu His Met Glu Ile Val Gly Thr Gln Gly Lys Pro	1022
305 310 315	
GCT ATT GCT CAC CGA GAT ATA AAG TCA AAG AAT ATC TTA GTC AAA AAG Ala Ile His Arg Asp Ile Lys Ser Lys Asn Ile Leu Val Lys Lys	1070
320 325 330	
TGT GAC ACT TGT GCC ATA GCT GAC TTA GGG CTG GCT GTG AAA CAT GAT Cys Asp Thr Cys Ala Ile Ala Asp Leu Gly Leu Val Lys His Asp	1118
335 340 345	
TCT ATC ATG AAC ACT ATA GAT ATA CCC CAG AAT CCT AAA GTG GGA ACC Ser Ile Met Asn Thr Ile Asp Ile Pro Gln Asn Pro Lys Val Gly Thr	1166
350 355 360 365	
AAG AGG TAT ATG GCT CCC GAA ATG CTT GAT GAT ACA ATG AAC GTC AAC Lys Arg Tyr Met Ala Pro Glu Met Leu Asp Asp Thr Met Asn Val Asn	1214
370 375 380	
ATC TTT GAG TCC TTC AAG CGA GCT GAC ATC TAT TCG GTG GGG CTG GTT Ile Phe Glu Ser Phe Lys Arg Ala Asp Ile Tyr Ser Val Gly Leu Val	1262
385 390 395	
TAC TGG GAA ATA GCT CGA AGG TGT TCA GTT GGA GGA CTT GTT GAA GAG Tyr Trp Glu Ile Ala Arg Arg Cys Ser Val Gly Gly Leu Val Glu Glu	1310
400 405 410	
TAC CAG TTG CCT TAT TAT GAC ATG GTG CCT TCA GAT CCT TCC ATA GAG Tyr Gln Leu Pro Tyr Tyr Asp Met Val Pro Ser Asp Pro Ser Ile Glu	1358
415 420 425	
GAA ATG AGG AAG GTC GTT TGT GAT CAG AAA CTG CGA CCA AAT CTC CCA Glu Met Arg Lys Val Val Cys Asp Gln Lys Leu Arg Pro Asn Leu Pro	1406
430 435 440 445	
AAC CAG TGG CAA AGC TGT GAG GCG CTC CGG GTC ATG GGA AGA ATA ATG Asn Gln Trp Gln Ser Cys Glu Ala Leu Arg Val Met Gly Arg Ile Met	1454
450 455 460	
CGT GAG TGC TGG TAT GCC AAC GGG GCA GCT CGC CTG ACC GCC CTG CGC Arg Glu Cys Trp Tyr Ala Asn Gly Ala Ala Arg Leu Thr Ala Leu Arg	1502
465 470 475	
GTG AAG AAG ACC ATT TCT CAG CTG TGT GTC AAG GAA GAC TGT AAG GCC Val Lys Lys Thr Ile Ser Gln Leu Cys Val Lys Glu Asp Cys Lys Ala	1550
480 485 490	
TAAGGATACA GGCGACGGGA AAGCCCTCAC CACTCTCTTT CATGTCTCCT GC	1602

Sub  
C1  
006TCT-15924.60

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 493 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEO ID NO:12:

Met 1	Thr	Pro	Ala	Arg 5	Arg	Ser	Ala	Leu	Ser 10	Leu	Ala	Leu	Leu	Val 15
Ala	Leu	Ala	Ser 20	Asp	Leu	Ala	Ala	Gly 25	Leu	Lys	Cys	Val	Cys 30	Leu
Cys	Asp	Ser 35	Ser	Asn	Phe	Thr	Cys 40	Gln	Thr	Glu	Gly 45	Ala	Cys	Trp
Ser 50	Val	Met	Leu	Thr	Asn 55	Gly	Lys	Glu	Gln	Val	Ser 60	Lys	Ser	Cys
Ser 65	Leu	Pro	Glu	Leu	Asn 70	Ala	Gln	Val	Phe	Cys 75	His	Ser	Ser	Asn
Val	Thr	Lys	Thr	Glu 85	Cys	Cys	Phe	Thr	Asp 90	Phe	Cys	Asn	Asn	Ile
Gln	His	Leu	Pro 100	Thr	Ala	Ser	Pro	Asp 105	Ala	Pro	Arg	Leu	Gly 110	Pro
Glu	Leu	Thr	Val	Val	Ile	Thr	Val 120	Pro	Val	Cys	Leu	Leu	Ser 125	Ile
Ala 130	Met	Leu	Thr	Ile	Trp 135	Ala	Cys	Gln	Asp	Arg	Gln 140	Cys	Thr	Tyr
Lys 145	Thr	Lys	Arg	His 150	Asn	Val	Glu	Glu	Pro	Leu 155	Ala	Glu	Tyr	Ser
Val	Asn	Ala	Gly 165	Lys	Thr	Leu	Lys	Asp 170	Leu	Ile	Tyr	Asp	Ala 175	Thr
Ser	Gly	Ser	Gly 180	Ser	Gly	Pro	Pro	Leu 185	Leu	Val	Gln	Arg	Thr 190	Ile
Arg	Thr	Ile 195	Val	Leu	Gln	Glu	Ile 200	Val	Gly	Lys	Gly	Arg 205	Phe	Gly
Val 210	Trp	His	Gly	Arg	Trp	Cys 215	Gly	Glu	Asp	Val	Ala 220	Val	Lys	Ile
Ser 225	Ser	Arg	Asp	Glu	Arg 230	Ser	Trp	Phe	Arg	Glu 235	Ala	Glu	Ile	Tyr
Thr	Val	Met	Leu	Arg 245	His	Glu	Asn	Ile 250	Leu	Gly	Phe	Ile	Ala 255	Ala
Asn	Lys	Asp	Asn 260	Gly	Thr	Trp	Thr	Gln 265	Leu	Trp	Leu	Val	Ser 270	Glu
His	Glu	Gln 275	Gly	Ser	Leu	Tyr	Asp 280	Tyr	Leu	Asn	Arg 285	Asn	Ile	Val
Val 290	Ala	Gly	Met	Val	Lys 295	Leu	Ala	Leu	Ser	Ile 300	Ala	Ser	Gly	Leu

His Leu His Met Glu Ile Val Gly Thr Gln Gly Lys Pro Ala Ile Ala  
305 310 315 320  
His Arg Asp Ile Lys Ser Lys Asn Ile Leu Val Lys Lys Cys Asp Thr  
325 330 335  
Cys Ala Ile Ala Asp Leu Gly Leu Ala Val Lys His Asp Ser Ile Met  
340 345 350  
Asn Thr Ile Asp Ile Pro Gln Asn Pro Lys Val Gly Thr Lys Arg Tyr  
355 360 365  
Met Ala Pro Glu Met Leu Asp Asp Thr Met Asn Val Asn Ile Phe Glu  
370 375 380  
Ser Phe Lys Arg Ala Asp Ile Tyr Ser Val Gly Leu Val Tyr Trp Glu  
385 390 395 400  
Ile Ala Arg Arg Cys Ser Val Gly Gly Leu Val Glu Glu Tyr Gln Leu  
405 410 415  
Pro Tyr Tyr Asp Met Val Pro Ser Asp Pro Ser Ile Glu Glu Met Arg  
420 425 430  
Lys Val Val Cys Asp Gln Lys Leu Arg Pro Asn Leu Pro Asn Gln Trp  
435 440 445  
Gln Ser Cys Glu Ala Leu Arg Val Met Gly Arg Ile Met Arg Glu Cys  
450 455 460  
Trp Tyr Ala Asn Gly Ala Ala Arg Leu Thr Ala Leu Arg Val Lys Lys  
465 470 475 480  
Thr Ile Ser Gln Leu Cys Val Lys Glu Asp Cys Lys Ala  
485 490

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGGATCCGT NGCNGTNAAR ATHTTYCC

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGGGATCCYT CNGGNGCCAT RTANCKYCTN GTNCC

00742564-121000

*Sub C1*